

A technique for lossy compression of error-diffused halftones

Sin-Ming Cheung and Yuk-Hee Chan

Centre for Multimedia Signal Processing
Dept of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

In this paper, a new technique for lossy compression of halftone images is proposed based on the vector quantization technique. A conventional vector quantization encoder is modified such that it embeds a block-based error diffusion process and takes a HVS model into account during the compression. This modification significantly improves the visual performance of encoded images while the compression ratio achieved is identical to that of vector quantization.

1. INTRODUCTION

Conventional binary image coding schemes such as JBIG [1] and G3 [2] are generally lossless and optimized for textual or graphical data instead of halftone images. In general, the achievable compression ratios of these schemes vary from 0.5 to 2.75[3] when halftones are handled. Lossy coding schemes are rare and they are usually dedicated for outputs of clustered dithering [4]. Their coding performance is not good for error-diffused halftones either.

In JBIG2[5], a lossy coding scheme for general halftones was recommended. Its high compression ratio is mainly achieved by the downsampling step in inverse halftoning. This step reduces the data bits required for representing the image, but it sacrifices the spatial resolution as well. In [6], Vallippan proposed an approach to reduce the amount of scarification in spatial resolution. His approach is composed of three stages including prefiltering, decimation, and quantization. At first, prefiltering is applied to reduce high frequency noise, spurious tone, and Nyquist frequencies. Then it reduces the spatial resolution by decimation and uses a modified error diffusion technique to shape the quantization error into the higher frequencies. This modification can improve the visual quality of the reconstructed halftone image without further reducing the spatial resolution as compared with the conventional approaches.

In this paper, we present a lossy compression scheme for compressing halftone images generated by error diffusion. In this proposed scheme, the simple vector quantization technique is used to compress the halftone images directly without downsampling. As a result, the spatial resolution can be preserved to a certain extent.

2. PROPOSED METHOD

Our method is based on the idea of vector quantization. Figure 1 shows the basic structure of a vector quantization codec. The input halftone image, S , is partitioned into a number of small non-overlapped blocks of size $M \times N$ each. For each block, say, $S_{x,y}$, we find a best-matched codeword of equal size, say, C_k , from a codebook $C = \{C_k : k = 0, 1, \dots, N_c\}$ to represent

it. The extent of matching is quantitatively determined with a predefined distortion function D . The corresponding codeword index k of the best-matched codeword is stored or transmitted. At the decoder side, an identical codebook is on hand and hence the decoder can restore the halftone image by performing table look-up.

Many different approaches can be used to generate codebooks for different interests. In this study, a codebook of clustered dot patterns as shown in Figure 2 was selected. These codebook favors printing applications since clustered patterns can resist ink spread. In this case, the codeword size is 4×4 and the size of the codebook, N_c , is 17.

2.1 Evaluation of the distortion

The encoder selects the codeword that provides the minimum distortion. The conventional mean square error (MSE) criterion used in vector quantization does not work when dealing with halftone images. In the implementation of our proposed scheme, a visual perception model is embedded in the distortion function D .

The spatial frequency sensitivity of our eyes is usually estimated as a modulation transfer function (MTF). According to [7], the impulse response of an one-dimensional eye filter to a printed image of 300 dpi at a viewing distance of 30 inches is virtually identical to that of a Gaussian filter of $\sigma = 1.5$ and $\tau = 0.0095^\circ$. These parameters are useful to define the one-dimensional best-fit model for a corresponding eye filter. In this research, the following 5×5 visual response filter H is used.

$$H = \frac{1}{11.566} \begin{bmatrix} 0.1628 & 0.3215 & 0.4035 & 0.3215 & 0.1628 \\ 0.3215 & 0.6352 & 0.7970 & 0.6352 & 0.3215 \\ 0.4035 & 0.7970 & 1 & 0.7970 & 0.4035 \\ 0.3215 & 0.6352 & 0.7970 & 0.6352 & 0.3215 \\ 0.1628 & 0.3215 & 0.4035 & 0.3215 & 0.1628 \end{bmatrix} \quad (3)$$

The input image is partitioned into a number of non-overlapped blocks and the blocks are encoded one by one in a raster scanning order. Let $S_{x,y}$ be a particular block and $O_{x,y}$ be the encoded version of $S_{x,y}$. In our case, each block contains 4×4 pixels and hence we have

$$S_{x,y} = \begin{bmatrix} S(4x, 4y) & S(4x, 4y+1) & \dots & S(4x, 4y+3) \\ S(4x+1, 4y) & S(4x+1, 4y+1) & & \vdots \\ \vdots & & \ddots & \vdots \\ S(4x+3, 4y) & \dots & \dots & S(4x+3, 4y+3) \end{bmatrix} \quad (4)$$

and

$$O_{x,y} = \begin{bmatrix} O(4x,4y) & O(4x,4y+1) & \cdots & O(4x,4y+3) \\ O(4x+1,4y) & O(4x+1,4y+1) & & \vdots \\ \vdots & & \ddots & \vdots \\ O(4x+3,4y) & \cdots & \cdots & O(4x+3,4y+3) \end{bmatrix} \quad (5)$$

where $S(i,j)$ and $O(i,j)$ are, respectively, the $(i,j)^{\text{th}}$ pixels of images S and O .

Without loss of generality, assume that we are now going to encode block $S_{x,y}$. The visual response of $\bar{S} = \begin{bmatrix} S_{x-1,y-1} & S_{x-1,y} \\ S_{x,y-1} & S_{x,y} \end{bmatrix}$

is estimated by $\hat{S} = \bar{S} \otimes H$, where \otimes denotes a two-dimensional convolution operator. This operation returns only those parts of the convolution that are computed without the zero-padded edges. It is expected that, after encoding, the mean square error between \hat{S} and the visual response of

$\bar{O} = \begin{bmatrix} O_{x-1,y-1} & O_{x-1,y} \\ O_{x,y-1} & O_{x,y} \end{bmatrix}$, which is estimated by $\hat{O} = \bar{O} \otimes H$, is minimum. That implies that one has to search the available codebook to get a codeword which minimizes the following objective function.

$$D = \|\hat{S} - \hat{O}_k\|^2 \quad (6)$$

where $\hat{O}_k = \bar{O}_k \otimes H$ is the estimated visual response of

$\bar{O}_k = \begin{bmatrix} O_{x-1,y-1} & O_{x-1,y} \\ O_{x,y-1} & C_k \end{bmatrix}$ and C_k is a particular codeword in the

codebook. Figure 3 shows the connection among \hat{O}_k , \bar{O}_k , C_k , \hat{S} and \bar{S} graphically.

2.2 Block-based error diffusion

Error diffusion is widely applied in digital halftoning as it can provide a very good visual quality of halftone at a reasonable cost. In our proposed compression scheme, error diffusion is performed in the encoder to change the nature of the noise introduced by vector quantization. Unlike halftoning, the diffusion is carried out block-wise as vector quantization is a block-based process.

Let $C_{x,y}$ be the optimal codeword for encoding block $S_{x,y}$ subject to the distortion criterion. The encoding output of $S_{x,y}$ is then given by $O_{x,y} = C_{x,y}$. The substitution of $S_{x,y}$ by $O_{x,y}$ introduces an error block in the estimated visual response domain. In formulation, we have

$$E_{x,y} = \hat{S} - \hat{O} \quad (7)$$

The effective area of $E_{x,y}$ covers 4×4 pixels as shown in Figure 4. The elements in $E_{x,y}$ correspond to the elements of E , an estimated visual response error plane associated with the output, as follows.

$$E_{x,y} = \begin{bmatrix} E(4x-2,4y-2) & E(4x-2,4y-1) & \cdots & E(4x-2,4y+1) \\ E(4x-1,4y-2) & E(4x-1,4y-1) & & \vdots \\ \vdots & & \ddots & \vdots \\ E(4x+1,4y-2) & \cdots & \cdots & E(4x+1,4y+1) \end{bmatrix} \quad (8)$$

The error in $E_{x,y}$ is then diffused to its neighboring blocks.

In practice, this can be achieved by diffusing the error of each element in $E_{x,y}$ to the neighboring pixels in a raster scanning order as shown in Figure 4. In our study, the error in $E_{x,y}$ is diffused to $\Omega_g = \{E(4x+2, j), E(i, 4y+2) \mid 4y-2 \leq j \leq 4y+2, 4x-2 \leq i \leq 4x+2\}$ only and the diffusion is carried out by diffusing elements in $E_{x,y}$ to Ω_g one by one as follows. Let $E(m,n)$ be a particular element in $E_{x,y}$. After its diffusion, the elements in $\Omega_{m,n,g} = \{(p,q) \mid p \geq m, q \geq n \text{ and } E(p,q) \in \Omega_g\}$ are updated as

$$E(p,q) = E(p,q) + E(m,n)w_{m,n,p,q}/W_{m,n} \quad \text{for } (p,q) \in \Omega_{m,n,g} \quad (9)$$

where $w_{m,n,p,q} = 1/\sqrt{(p-m)^2 + (q-n)^2}$ for $(p,q) \in \Omega_{m,n,g}$

and $W_{m,n} = \sum_{(p,q) \in \Omega_{m,n,g}} w_{m,n,p,q}$. After diffusing all error in $E_{x,y}$, all elements of $E_{x,y}$ are set to be zero.

The error diffused to the neighboring blocks should be taken into account when these neighboring blocks are encoded later on. To achieve this, the objective function in eqn.(6) is modified as follows.

$$D = \|\hat{S} + \hat{E} - \hat{O}_k\|^2 \quad (10)$$

where $\hat{E} = \begin{bmatrix} E(4x,4y) & E(4x,4y+1) & \cdots & E(4x,4y+3) \\ E(4x+1,4y) & E(4x+1,4y+1) & & \vdots \\ \vdots & & \ddots & \vdots \\ E(4x+3,4y) & \cdots & \cdots & E(4x+3,4y+3) \end{bmatrix}$ is

the error diffused into the effective area of $E_{x,y}$ earlier on. The initial values of the elements of E are all zero.

Similarly, after selecting a codeword for a particular block, $E_{x,y}$ should include both the error introduced by the substitution of the codeword and the components diffused from previously encoded blocks. Hence, when block-based error diffusion is realized during the encoding, eqn.(7) should be modified to be

$$E_{x,y} = \hat{S} + \hat{E} - \hat{O} \quad (11)$$

3. SIMULATION RESULT

Simulations were carried out to evaluate the performance of the proposed scheme with a set of halftone images of size 256×256 . The testing halftone images were generated with standard error diffusion [8]. The HVS filter was on and the block-based error diffusion filter was purposely on or off in different settings to evaluate their effect on the coding performance. Some other schemes including JBIG1 [1] and Vallippan's scheme [6] were also simulated for comparison.

Table 1 shows the compression ratios of the schemes at different settings. In Table 1, one can see that the lossy compression schemes can provide a 2- to 3-fold extra compression gain as compared with JBIG1. Note that JBIG1 is a lossless compression scheme. Vallippan's scheme is a VQ-like block-based coding scheme for encoding halftone images [6]

and hence its performance was also evaluated for comparison. To obtain the figures of Vallippan's and our schemes in Table 1, their codeword index maps were treated as 17-level gray-level images and further encoded with JPEG-LS. To improve the coding performance, codebooks were reindexed with the technique proposed in [9] if necessary. On the average, the compression ratio of the proposed scheme is lower as compared with Vallippan's scheme. JPEG-LS is not suitable for encoding the index maps generated with the proposed scheme.

Figure 5 shows some simulation results for subjective evaluation. In general, the compression results of Vallippan's approach are blurred while the proposed approach can provide visually sharper results after compression. In other words, more feature details such as edges can be preserved with the proposed scheme. The reduction in the spatial resolution is apparently less and the contrast is apparently higher. One can see the difference by examining the tablecloth, the face, books in the bookshelf, the scarf and the object on the table. Note that the two outputs are produced with the same codebook.

An objective quality measure called Hierarchical Intensity Distribution (HID) was proposed by Katsavounidis to describe how close a halftone image is to its original gray-level image in different scales [10]. This measure can be used here to evaluate the closeness of an original halftone and its encoded output.

A digital halftone image uses a group of pixels to emulate a gray level. Accordingly, the effective gray level of a pixel at a particular location can be defined as the total number of white dots in a small local region centered at the pixel. In HID, the distortion at a particular scale is defined as

$$Q = \sum_{\forall m,n} \left(\sum_{(k,l) \in R_{m,n}} S(k,l) - \sum_{(k,l) \in R_{m,n}} O(k,l) \right)^2 \quad (12)$$

where $R_{m,n}$ defines the local region of the pixel at location (m,n) , and, $S(k,l)$ and $O(k,l)$ are, respectively, values of corresponding pixels at the original halftone image and the encoded halftone image. In this research, $R_{m,n}$ is a square window of size $N_w \times N_w$, where N_w can be any integer. The values of Q obtained for different N_w tell the closeness of the images at different scales, which are plotted in Figure 6.

The following observations can be obtained from these Figures. First, in terms of HID, the proposed scheme is superior to Vallippan's scheme. Second, the Q value achieved with the codebook of cluster dot patterns is remarkably low when N_w is a multiple of 4. This is because the codewords are 4×4 cluster dot patterns which favors the measure at corresponding scales.

4. CONCLUSIONS

In this paper, we propose a new technique for lossy compression of halftone images. We modify a conventional vector quantization encoder such that it embeds a block-based error diffusion process and takes a HVS model into account during the compression. This modification improves the visual performance of encoded images while maintaining the same number of codeword indices as vector quantization does at the output of the compression.

As compared with Vallippan's scheme, simulation results show that the proposed scheme can provide a subjectively better encoding output with the same codebook. It is better in a way that the reconstructed images are sharper and more features and detail can be preserved in the images. Evaluation results also show that the proposed scheme is superior to Vallippan's scheme in terms of Hierarchical Intensity Distribution. However, the compression ratio of the proposed scheme is lower when JPEG-LS is used to encode the codeword indices.

This paper presents some preliminary results obtained in our study of the proposed scheme. Factors such as the codebook exploited, the block-based error diffusion filter and the lossless encoder used for encoding the codeword indices can affect the rate-distortion performance significantly and should be further explored in the future.

5. ACKNOWLEDGEMENTS

This work was substantially supported by a grant from the Centre for Multimedia Signal Processing, HKPolyU (A046).

REFERENCES

- [1] International Organization for Standards/International Electrotechnical Commission (ISO/IEC), "Progressive Bi-level Image Compression," Int. Standard 11544: 1993
- [2] International Telegraph and Telephone Consultative Committee (CCITT), "Standardization of Group 3 Facsimile Apparatus for Document Transmission," Recommendation T.4, 1980
- [3] R.B. Arps and T.K. Truong, "Comparison of international standards for lossless still image compression," *Proc. IEEE*, vol.82, June 1994, pp. 889-899
- [4] R.A.V. Kam, R.M. Gray, "Lossy compression of clustered-dot halftones using sub-cell prediction" *Proc., Data Compression Conference*, 1995, pp. 112-121
- [5] JBIG Committee, ISO/IEEE JTC1/SC29/WG1 (ITU-T SG8) WD14492, July 1999
- [6] M. Valliappan, B.L. Evans, D.A.D. Tompkins and F. Kossentini, "Lossy compression of stochastic halftones with JBIG2," *Proc., IEEE ICIP'99*, Vol.1, Dec 1999, pp.214-218
- [7] T.N. Pappas, D. L. Neuhoff, "Least-squares model-based halftoning," *Proc., SPIE*, Human Vision, Visual Processing and Digital Display III, Vol.1666, 1992, pp.165-176
- [8] R.W. Floyd and L. Steinberg, "An adaptive Algorithm for Spatial Gray Scale," *Proceedings of SID International Symposium Digest of Technical Papers*, Society for Information Displays, 1975, pp.36-37
- [9] A. J. Pinho and António J. R. Neves, "A note on Zeng's technique for color reindexing of palette-based images," *IEEE Signal Processing Letters*, Vol.11, No.2, Feb 2004, pp.232-234.
- [10] I. Katsavounidis and C.C. Kuo, "A multiscale error diffusion technique for digital halftoning," *IEEE trans. on image processing*, Vol.6, Mar 1997, pp.483-490

	JBIG I [1]	Vallippan's [6]	Ours	
			w/o ED	ED
Peppers	1.635	6.942	5.076	5.041
Lenna	1.617	7.167	5.531	5.393
Baboon	1.157	7.111	4.917	4.900
Airplane	1.716	7.288	4.850	4.876
House	2.033	9.092	6.671	6.533
Bridge	1.581	7.099	5.010	4.986
Cameraman	1.810	8.135	5.592	5.411
Girl	1.983	7.907	6.596	6.554
Goldhill	1.625	7.961	5.418	5.440
Barbara	1.527	7.111	5.172	5.095
Ship	1.498	7.407	5.313	5.299
Average	1.653	7.566	5.468	5.412

Table 1 Compression ratios achieved by different methods

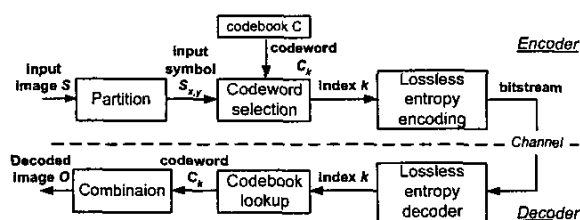


Figure 1 Block diagram of a VQ codec

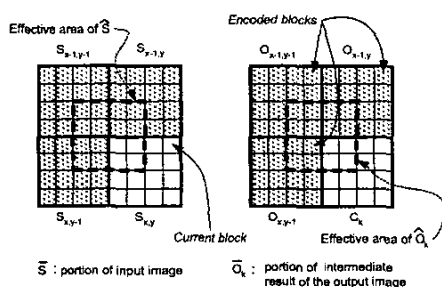


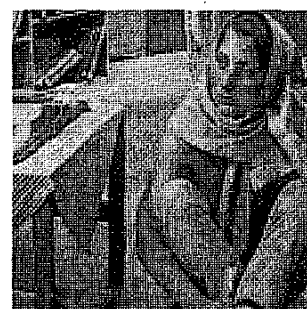
Figure 3 Template for evaluating the distortion between the input and the candidate constructed with a codeword.



Original



Vallippan's [6]



Ours (with ED)

Figure 5 Performance of various lossy coding schemes with a codebook of cluster dot patterns.

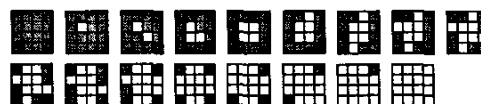


Figure 2 A codebook of cluster patterns

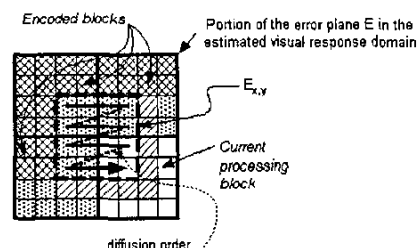


Figure 4 Block-based error diffusion

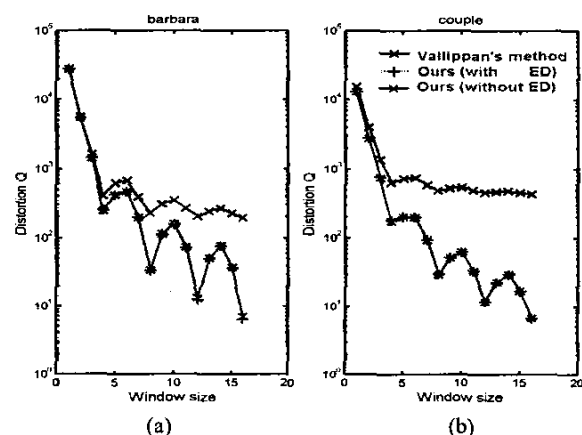


Figure 6 Performance of various coding schemes in terms of Hierarchical Intensity Distribution with a codebook of cluster dot patterns. (a) "Barbara" and (b) "Couple".